[0078] The user identification unit may be configured to provide an identifier to the object pose detector **203** indicating which user a given object corresponds to. The object pose detector **203** may provide the identifier to the user input generator **204**, such that changes in pose of a given object are used to generate user inputs for the corresponding user.

[0079] In some examples, the video game may provide some visual feedback as to the objects that have been detected and assigned to respective users. For example, if the video images of the users holding the objects are displayed at a display, then a visual indication of 'player **1**' and 'player **2**' may provided at positions corresponding to the respective positions of the players in the images. This may allow users to immediately identify which character or virtual object they are controlling.

[0080] The user input generator **204** may be configured to generate inputs for each user participating in the video game session based on the changes in pose of the object identified as corresponding with that user. The detection and tracking of the objects being held by the user may be performed in any of the previously described manners. In some examples, each user may hold more than one object, with the detection and tracking occurring as described previously. The user input generator **204** may be configured to transmit an identifier associated with each detected object (or e.g. object pair) to the CPU of the video game playing device, to ensure that the generated user inputs are interpreted as belong to different users.

[0081] FIG. **7** schematically illustrates a modified version of the system **700** shown in FIG. **2**. In FIG. **7**, the system **700** is adapted so as to generate an augmented image of the object or objects being held by the user. The system **700** comprises an input unit **701**, object detector **702**, object pose detector **703** and user input generator **704**; each of these components may operate in the manners described previously.

[0082] In FIG. **7**, the system **700** further comprises an image generator **705** operable to generate an image of one or more virtual buttons for superimposing on top of a user's view of the non-luminous object. The image generator **705** may be located at e.g. the video game playing device, or a peripheral device such as an HMD.

[0083] In FIG. **7**, the system **700** also comprises a display **706** for displaying the image of the virtual button at a location that corresponds to a location on the surface of the non-luminous object. The display **706** may correspond to the display of a head-mountable display (HMD), with the image of the virtual button corresponding to an augmented reality image. The HMD may be configured to obtain information indicating a relative position and depth of the object, and to display the image of the virtual button at a location on the display that coincides with a viewable surface of the object. The position and depth of the object relative to the user may be obtained from a camera, such as the camera that captured the images of the user holding the object, for example.

[0084] Alternatively, or in addition, the display **706** may correspond to e.g. a TV device that is being used to play the video game, and at which the video image of the user holding the object is visible. In such examples, the image of the virtual button(s) may be displayed at a location in the video image of the user that corresponds to a surface of the object(s) being held by the users. In such examples, the location of the object in the images may be detected as described previously (e.g. via machine learning and/or contour detection), and the images of the virtual buttons displayed at a location that is within the pixel area corresponding to the detected object(s).

[0085] In FIG. **7**, the system **700** also comprises a finger detector **707** operable to detect a user's finger in the obtained images and a location of the finger relative to the non-luminous object. The finger detector **707** is configured to detect when the location of a user's finger relative to the object coincides with the location of the virtual button. The finger detector **707** is configured to provide an input to the user input generator **704** in response to detecting that a user's finger intersects a (or one of) the virtual buttons. The user input generator **704** is configured to generate a user input in accordance with a pressing of the virtual button. The user input generated in response to the pressing of the virtual button may correspond to a further, additional input that is different from the inputs generated in response to detected movement of the non-luminous object.

[0086] The finger detector **707** may be configured to detect the fingers of the user in the obtained images in a similar manner to the detection of the objects themselves. For example, the finger detector **707** may comprise a machine learning model that has been trained to identify both objects and fingers holding those objects in images. Alternatively, or in addition, contour detection may be used to detect the fingers in the obtained images; for example, if a first contour is detected as corresponding to a given object, contours located between or adjacent to that contour, and having a relative size that is comparable to the relative size of fingers, may be identified as such.

[0087] In FIG. **7**, the finger detector **707** is shown as receiving an optional input from the object detector **702**. The input from the object detector **702** may include e.g. an indication of the locations of the fingers relative to the objects detected in the obtained images. That is, the object detector **702** may be configured to detect the fingers and the objects in the obtained images and to provide an indication of this to the finger detector **707**, which then determines whether a user is attempting to interact with the virtual button. In other examples, it may be that the finger detection is performed separately, using e.g. a separate camera secured to the front of an HMD.

[0088] By superimposing virtual buttons on top of a player's view of the non-luminous object, further granularity can be added to the controller.

[0089] FIG. **8** illustrates schematically an example of how the non-luminous object may appear to the user when one or more virtual buttons have been superimposed on top of the user's view **800** of the object. In FIG. **8**, it can be seen that the banana **802**, being held by the user **804**, has two virtual buttons **806** displayed at locations correspond to the user's view of the front surface (relative to the viewer) of the banana. The image **800** shown in FIG. **8** may correspond to the image displayed at an HMD (either as a video image of the user) or as an image overlaid on top of a transparent display through which the user looks. Alternatively, image **800** may correspond to a video that is displayed at the same display device as the display being used to play the video game.

[0090] FIG. **9** shows schematically an example of a method for generating video game inputs in accordance with the present disclosure. The method corresponds to the steps performed by the system described previously. As above, it